Coding Opportunity Densification Strategies for Instantly Decodable Network Coding

Sameh Sorour, *Member, IEEE,* Shahrokh Valaee, *Senior Member, IEEE*

Abstract

In this paper, we identify the strategies that maximize the density of the coding opportunities in instantly decodable network coding (IDNC). Using the well-known graph representation of IDNC, we derive expressions for the exact and expected evolutions of coding opportunities after the transmission of any arbitrary coded packet. From the exact expression, we show that serving packets requested by a larger number of receivers tends to maximize the number of coding opportunities. We then employ the expected expression to show that serving the maximum number of receivers, having the largest numbers of missing packets and erasure probabilities, tends to both maximize and monotonically increase the expected density of coding opportunities. Simulation results justify our theoretical findings. Finally, we validate the importance of our work through a case study showing that our identified strategy to increase the coding density outperforms the step-by-step service maximization solution in reducing the IDNC completion delay.

Index Terms

Instantly Decodable Network Coding; Coding Opportunities; Wireless Broadcast; Graph Densification.

I. Introduction

Network coding (NC) [1] has shown great abilities to substantially improve transmission efficiency, throughput and delay over broadcast erasure channels. The design of network coding algorithms, optimizing throughput and delay performances over single-hop broadcast erasure channels, has recently been an intensive area of research [2]–[11]. Some of these works have focused on packet selection for coding in each transmission in order to optimize a certain metric, such as in-order delay [11] or video quality [9], [10]. Other works have focused more on receiver selection in each transmission to optimize another

The authors are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, M5S 3G4, Canada, e-mail:{samehsorour, valaee}@comm.utoronto.ca.

set of parameters, such as completion and decoding delay [7], [12]–[14]. These works have considered a subclass of network coding known as *instantly decodable network coding (IDNC)*, in which coded packets must be decoded at their reception instant and cannot be stored for future decoding. This opportunistic network coding scheme has attracted attention due to its desirable properties, such as fast packet recovery, simple XOR coding and decoding, and no buffer requirements.

In most of these opportunistic network coding and IDNC works, the selection of targeted receivers and combined packets, to optimize a desired parameter for a particular transmission, does not consider the effect of this selection on resulting *"coding opportunities"* in subsequent transmissions. By a coding opportunity, we mean the opportunity of serving two packet requests of two receivers simultaneously by one transmission using network coding. For instance, the proposed online algorithms in [3], [4], [7], [12], [14] have focused on increasing the number of decoding receivers in each transmission, without studying the effect of such approach on the number of remaining coding opportunities for subsequent transmissions. If this selection leads to a very limited number of opportunities, then the sender will no longer be able to send packets that are decodable by many receivers, thus sacrificing the main strength of network coding and its overall performance. Instead, if these algorithms consider the maintenance of a larger number of coding opportunities for subsequent transmissions, they may end-up with an overall better performance. In fact, it can be inferred from [13], [14] (from eq. (7) in both references) that maintaining a large number of coding opportunities in the *"IDNC graph"* plays a role in optimizing the IDNC completion and decoding delays along the course of the transmission of a frame of packets.

To give an example of this role, let us assume a network of 6 receivers $\{r_1, \ldots, r_6\}$ that require packets $\{p_1, \ldots, p_6\}$, respectively. Also assume that, due to the side information at the different receivers, the available coding opportunities are $\{p_1 \oplus p_2 \oplus p_3, p_1 \oplus p_4, p_2 \oplus p_5, p_3 \oplus p_6\}$. Intuitively, one may think that the optimal completion and decoding delays will be obtained if the maximum number of packet requests are served in each transmission (similar to the algorithms in [3], [4], [7], [12], [14]). In this case, the sender should first send the packet combination $p_1 \oplus p_2 \oplus p_3$ as it serves the maximum number of requests in the first transmission. However, this selection will result in three packets $\{p_4, p_5, p_6\}$ with no coding opportunities between them, and thus require 3 additional transmissions to be served. Moreover, r_4, r_5 and r_6 will not be able to decode any packets until the second, third and fourth transmissions, respectively, and thus will suffer from decoding delays of 1, 2 and 3 time units, respectively. On the other hand, a first transmission serving a smaller number of packet requests (such as $p_1 \oplus p_4$) keeps 2

coding opportunities in the system, which can be satisfied by only 2 additional transmissions. Moreover, the decoding delay that could be experienced by any receiver in this case is either 1 or 2. Consequently, the selection of coded transmissions that preserve a large number of coding opportunities in the system results in a better completion and decoding delays.

Despite its importance illustrated by the above example, the number of coding opportunities may not be expressive in itself. Indeed, a selected transmission may result in a larger number of coding opportunities because it serves very few packet requests, and thus the number of remaining requests will be also large. Thus, this apparently large number of coding opportunities will not be enough to foster efficient combinations between the large number of remaining requests. Consequently, it is important not only to maximize the absolute number of coding opportunities but to mainly maximize their ratio to the number of remaining packet requests. We thus define the *coding density* as the number of actual coding opportunities normalized by the maximum number of coding opportunities that could exist for the same number of packet requests. Consequently, this coding density parameter evaluates the number of coding opportunities with respect to the total number of these requests.

In this paper, we aim to answer the following question: *What are the receiver and packet selection strategies in IDNC that can maximize the coding density after each transmission and over the transmission horizon of a frame of packets?* To answer this question, we first review the graph representation of IDNC, in which the edge set of the graph represents the set of all coding opportunities between all missing packet requests by all receivers. We then derive an expression for the exact evolution of the edge set size after the transmission of any arbitrary coded packet. From the derived expressions, we show that an efficient strategy for increasing the number of coding opportunities is to select packets that are wanted by the maximum number of receivers, especially in the beginning of the packet recovery phase. Since the exact evolution expression does not give clear insights on the receiver selection strategy, we derive the expectation of the edge set size evolution after ignoring the identities of the packets requested by the different receivers and considering only their numbers. From this expression, we show that the best strategy to increase the expected coding density is to serve the maximum number of receivers having the largest number of missing packets and erasure probabilities. We then test our identified strategies and compare them with other well-known IDNC strategies. Finally, we validate the importance of our study and the chosen metrics by presenting a case study showing the effect of the identified receiver selection strategies on reducing the completion delay in IDNC.

It is important to note that this paper is not proposing algorithms to optimize any specific throughput or delay parameter. It is rather a first and independent study of an influential component in network coding, namely the coding density, which has been totally ignored in most works on designing network coding algorithms, despite its clear importance in optimizing long-term parameters compared to per-transmission benefits. The contributions of the paper can be summarized as follows:

- It derives expressions for the evolution of coding opportunities and density along the transmission of network coded packets.
- It provides a rigorous analysis of the parameters affecting the evolution of coding opportunities and density.

These contributions can help, guide and open paths for future efforts in designing more efficient online network coding algorithms, optimizing different performance metrics, when our coding density analysis is incorporated in the study of these metrics. The paper finally illustrates the importance of coding density by a case study on completion delay.

The rest of the paper is organized as follows. In Section II, the system model and parameters are illustrated. We introduce the IDNC graph and our metric of coding density in Section III. In Section IV, we derive the expression for the exact edge set size evolution and identify, in Section V, the packet selection strategy increasing it. We then derive the expected edge set size evolution in Section VI and identify, in Section VII, the receiver selection strategy increasing it. Simulation results are illustrated in Section VIII. Section IX presents a case study of the IDNC completion delay optimization in the light of the paper's findings. Finally, Section X concludes the paper.

II. System Model and Parameters

The model consists of a wireless sender that is required to deliver a frame (denoted by $\mathcal{N}$) of N source packets to a set (denoted by $\mathcal{M}$) of M receivers. The sender initially transmits the N packets uncoded in an *initial transmission phase*. Each sent packet can be successfully received at receiver i with probability q_i, which is assumed to be fixed during the frame transmission period. Receivers feed back to the sender a positive one-bit acknowledgement (ACK) for each received packet. Consequently, an overhead of $O(N)$ bits is required for feedback after each transmission. At the end of the initial transmission phase, two sets of packets are attributed to each receiver i, representing the feedback state of the network:

- The *Has* set (denoted by $\mathcal{H}_i$) is defined as the set of packets correctly received by receiver i.

- The *Wants* set (denoted by $\mathcal{W}_i$) is defined as the set of packets that are not yet received by receiver i. In other words, $\mathcal{W}_i = \mathcal{N} \setminus \mathcal{H}_i$.

The cardinalities of the Has and Wants sets of receiver i are denoted by ϱ_i and ψ_i, respectively. After the initial transmission phase, a recovery phase starts, in which the sender exploits the diversity of received packets to transmit network coded combinations. According to the definition of IDNC, these combinations must be either decoded at their reception instant or discarded. The received ACKs at the sender after each transmission are used to update the different sets. This process is repeated until all receivers obtain all the packets.

III. The IDNC Graph

The IDNC graph defines the set of all feasible instantly decodable packet combinations. It was first introduced in the context of a heuristic algorithm design solving the index coding problem [15]. The IDNC graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is constructed by first generating a vertex v_{ij} in $\mathcal{V}$ for each packet $j \in \mathcal{W}_i, \forall\, i \in \mathcal{M}$. Two vertices v_{ij} and v_{kl} in $\mathcal{G}$ are adjacent if one of the following conditions is true:

- C1: $j = l \Rightarrow$ The two vertices represent the loss of the same packet j by the two receivers i and k.
- C2: $j \in \mathcal{H}_k$ and $l \in \mathcal{H}_i \Rightarrow$ The requested packet of each vertex is in the Has set of the receiver that induced the other vertex.

Consequently, each edge between two vertices in the graph represents a coding opportunity that is instantly decodable for receivers inducing these vertices. Given this graph formulation, we can easily define the set of all feasible packet combinations in IDNC as the set of packet combinations defined by all cliques in $\mathcal{G}$. Consequently, the sender can generate an IDNC packet for a given transmission by XORing all the packets identified by the vertices of a selected clique κ in $\mathcal{G}$. According to the design of $\mathcal{G}$, we can easily infer that any clique κ in $\mathcal{G}$ can include at most one vertex induced by any given receiver to maintain instant decodability. In the rest of the paper, we say that an IDNC packet *targets* a receiver if the corresponding clique includes a vertex belonging to this receiver.

Based on this modeling of coding opportunities as the edges of a graph, we can define the coding opportunity density (or coding density for short) $\rho_c(\mathcal{G})$ for IDNC as the density of its graph $\mathcal{G}$. In graph theory, the graph density is the ratio of the total number of edges in this graph to the number of edges of a complete graph with the same number of vertices. We can express this graph density (and thus coding

density) as:

$$\rho_c(\mathcal{G}) = \frac{|\mathcal{E}|}{\frac{1}{2}|\mathcal{V}|(|\mathcal{V}|-1)} \tag{1}$$

It is obvious that the maximization of $\rho_c(\mathcal{G})$ guarantees a large number of coding opportunities with respect to the number of remaining packet requests (i.e. vertices), and thus a large number of receivers and packet requests can be served simultaneously in each IDNC packet.

From the above expression, we can see that, in order to maximize the coding density in each step, the selected cliques should be able to both maximize the number of edges and minimize the vertex set size. The number of vertices is clearly minimized by serving the maximum number of receivers in each transmission. However, this selection may decrease the coding density if the numerator is significantly reduced. To study this effect, we first need to derive an expression for the edge set size evolution after any arbitrary transmission.

IV. Exact Coding Opportunity Evolution

In order to derive an expression for the edge set size evolution, we start by deriving the expression of the edge set size for any given feedback state.

A. Edge Set Size

Theorem 1 introduces the expression of the edge set size of the IDNC graph.

Theorem 1. *The edge set size for an arbitrary feedback state can be expressed as:*

$$|\mathcal{E}| = \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M}(\psi_{ik}+\theta_{ik}\theta_{ki}), \qquad \textit{where} \quad \psi_{ik}=|\mathcal{W}_i\cap\mathcal{W}_k|, \quad \theta_{ik}=\psi_i-\psi_{ik}, \quad \theta_{ki}=\psi_k-\psi_{ik}\,. \tag{2}$$

Proof: The proof can be found in Appendix A. ■

The intuition behind Theorem 1 is illustrated in Figure 1. The figure depicts the number of pairwise edges between receivers i and k (denoted by Y_{ik}). As shown, the vertices of i and k can be classified into two sets. The gray vertices represent the vertices of i and k requesting the same packets (i.e. pairs of vertices of i and k with $j=l$). Thus, such vertex pairs are adjacent according to condition C1 in Section III. They cannot be adjacent to other vertices of the opposite receiver as this will violate condition C2. We say that each of these vertices is pairwise restricted by its adjacent vertex at the opposite receiver. Thus, these vertices will contribute to Y_{ik} by $|\mathcal{W}_k\cap\mathcal{W}_i|=\psi_{ik}$ edges as shown in Figure 1.

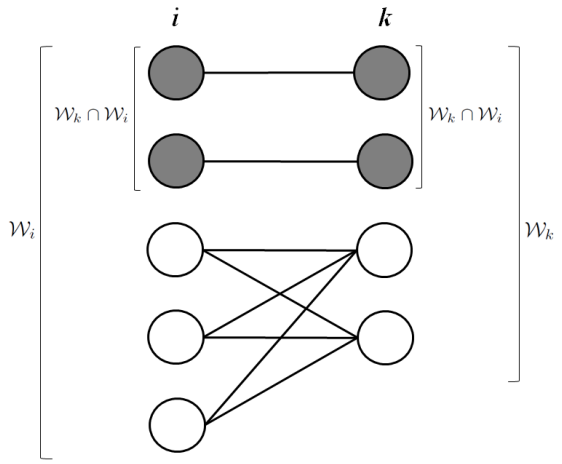

Fig. 1. Pairwise edges between receivers i and k given their Wants sets. The grey vertices represent the packets that are commonly wanted by both receivers whereas the white vertices represent packets that are wanted by either of them and has been received by the other.

The white vertices represent the mutually unrestricted vertices (i.e. not in $\mathcal{W}_k \cap \mathcal{W}_i$), and can all be connected to each other as a full bipartite subgraph because they all satisfy condition C2. Consequently, they contribute to Y_{ik} by $(\psi_i - \psi_{ik})(\psi_k - \psi_{ik})$ edges. The final expression in (2) results from summing the Y_{ik}'s of all i and $k \neq i$ and dividing by half to remove repetitions.

B. Edge Set Size Evolution

Before deriving the expression of the edge set size evolution, we will first illustrate the different possibilities of evolution on the pairwise subgraph in Figure 1, when either or both receivers i and k are targeted with source packets p_i and p_k, respectively, in one IDNC transmission. These possibilities are illustrated in Figure 2, in which all preserved edges from $Y_{ik}^{(t)}$ to $Y_{ik}^{(t+1)}$ are removed for ease of illustration. The served vertices are marked in black and added (removed) edges are represented by solid (dashed) lines.

In Figure 2(a), only one receiver (in this case k) is targeted with an unrestricted vertex v_{kp_k} with respect to the other receiver (in this case i). If this receiver receives, its vertex v_{kp_k} and all its adjacent edges (which are shown in Figure 1) will disappear from the graph. Since the opposite vertices to these removed edges are already unrestricted with respect to k, they will not gain any additional edges.

In Figure 2(b), only one receiver (in this case k) is targeted with a restricted vertex v_{kp_k} by the other receiver (in this case i). If this receiver k receives, its vertex v_{kp_k} will disappear as well as its edge to the corresponding restricted vertex of i (i.e. v_{ip_k}). Consequently, this restricted vertex v_{ip_k} becomes unrestricted with respect to k and thus becomes adjacent to all k's unrestricted vertices with respect to i.

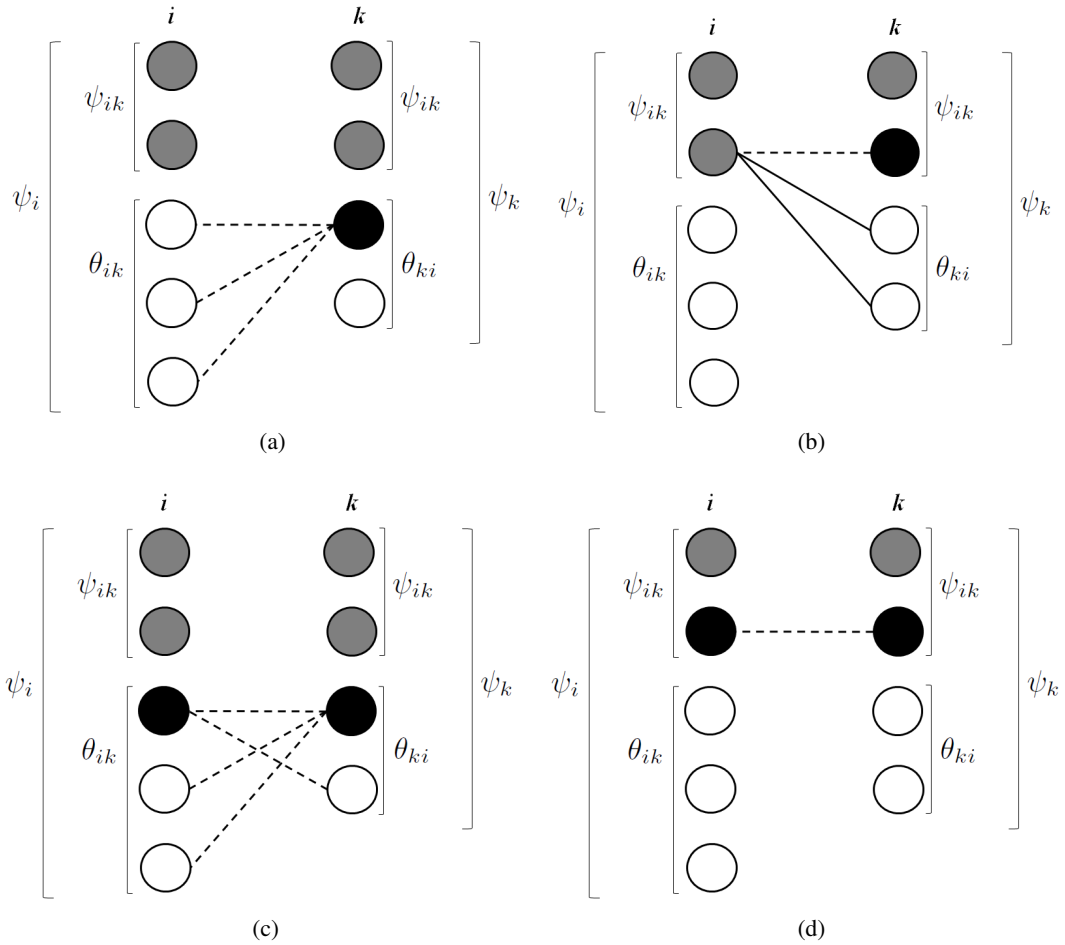

Fig. 2. Added (solid) and removed (dashed) edges between the vertices of receivers i and k in case: (a) Only one of the two receivers is targeted with an unrestricted vertex. (b) Only one receiver is targeted with a restricted vertex. (c) Each of the two receivers is targeted with an unrestricted vertex and they both receive. (d) Each of the two receivers is targeted with a restricted vertex and they both receive.

The third possibility is when both receivers are targeted with one of their mutually unrestricted vertices. If only one of them receives, we get the case of Figure 2(a). If both receivers receive, each of the vertices will behave as in Figure 2(a) resulting in the evolution in Figure 2(c).

The last possibility is when both receivers are targeted with one of their mutually restricted vertices. If only one of them receives, we get the case of Figure 2(b). If both receivers receive, they both disappear with their common edge and without addition of new mutual edges, as shown in Figure 2(d).

After the description of these possibilities, we can now introduce the following theorem.

Theorem 2. *For an arbitrary attempted clique κ at time t with a set of targeted receivers $\mathcal{T}$, the edge*

set size at time $t+1$ after this attempt can be expressed as:

$$\left|\mathcal{E}^{(t+1)}\right| = \left|\mathcal{E}^{(t)}\right| + \frac{1}{2}\sum_{\substack{i\notin\mathcal{T},k\in\mathcal{T}\\ p_k\in\mathcal{W}_i}} X_k\hat{\theta}_{ki} - \frac{1}{2}\sum_{\substack{i\notin\mathcal{T},k\in\mathcal{T}\\ p_k\notin\mathcal{W}_i}} X_k\theta_{ik} + \frac{1}{2}\sum_{\substack{i\in\mathcal{T},k\notin\mathcal{T}\\ p_i\in\mathcal{W}_k}} X_i\hat{\theta}_{ik} - \frac{1}{2}\sum_{\substack{i\in\mathcal{T},k\notin\mathcal{T}\\ p_i\notin\mathcal{W}_k}} X_i\theta_{ki}$$
$$- \frac{1}{2}\sum_{\substack{\{i,k\}\in\mathcal{T}\\ p_i\notin\mathcal{W}_k}} \left(X_i\theta_{ki} + X_k\theta_{ik} - X_iX_k\right) + \frac{1}{2}\sum_{\substack{\{i,k\}\in\mathcal{T}\\ p_i\in\mathcal{W}_k}} \left(X_i\hat{\theta}_{ik} + X_k\hat{\theta}_{ki} - X_iX_k\left(\hat{\theta}_{ik} + \hat{\theta}_{ki} - X_iX_k\right)\right), \quad (3)$$

where $\hat{\theta}_{ik} = \theta_{ik} - 1$, $\hat{\theta}_{ki} = \theta_{ki} - 1$, and X_h is the reception indicator of receiver h, which is equal to 1 if h receives the transmitted packet and zero otherwise.

Proof: The proof can be found in Appendix B. ■

Using this formula, we can investigate the packet selection strategy in the next section.

V. Packet Selection Strategy

From (3), we can prove the following theorem for any two arbitrary receivers.

Theorem 3. *Targeting either one or both receivers i and k with a packet in $\mathcal{W}_i \cap \mathcal{W}_k$ (i.e. a common wanted packet) at time t results in a greater or equal number of pairwise edges between them at time $t+1$ $\left(\text{i.e. } Y_{ik}^{(t+1)}\right)$ compared to targeting either one or both receivers with packets that are not in $\mathcal{W}_i \cap \mathcal{W}_k$. .*

Proof: The proof can be found in Appendix C. ■

The above theorem proves that serving a common wanted packet for receivers i and k always results in a larger increase or smaller reduction in the number of their pairwise edges $Y_{ik}^{(t+1)}$, whether both receivers are targeted with this packet or only one of them is. Now, since Equation (2) in Theorem 1 expresses the overall edge set size as a linear addition of these numbers of pairwise edges, then having this property satisfied for all receivers (i.e. the sender transmits a packet that is in the Wants set of all receivers at time t) will result in the maximum edge set size that could be achieved at time $t+1$ according to Theorems 1 and 3. Violating this condition for any one receiver i will replace its $Y_{ik}^{(t+1)}\ \forall\ k \in \mathcal{M} \setminus i$ in (2) with smaller values, which will lead to smaller edge set size. Increasing the number of receivers violating this condition will result in further replacements with smaller terms, and thus a larger reduction in the edge set size.

Thus, we can conclude that serving the packets that are wanted by the maximum number of receivers tends to maximize the resulting edge set size and thus the number of coding opportunities. We will refer to this strategy as the *Most Wanted Packet Serving (MoWPS) strategy*. We can express this strategy as

choosing the maximal clique κ^* in each transmission such that:

$$\kappa^* = \arg\min_{\kappa \in \mathcal{G}} \sum_{j|v_{ij} \in \kappa} |\mathbf{\Omega}_j|^n \qquad \text{s.t.} \quad \mathbf{\Omega}_j = \{i \in \mathcal{M} | j \in \mathcal{W}_i\} \ , \tag{4}$$

where $\mathbf{\Omega}_j$ is the set of receivers wanting packet j and n is a biasing factor. In other words, the MoWPS strategy selects the maximum weight clique in the IDNC, such that the weight of each vertex v_{ij} in the graph is defined by $\mathbf{\Omega}_j$. Thus, this maximum weight clique will include the vertices representing the packets that are most wanted by the receivers.

However, this finding does not mean that the MoWPS strategy will always increase the coding density at all stages of the recovery transmission phase. The increase of coding density, after any transmission, will be conditioned on the presence of large sets of receivers requesting the same packets for this transmission. When this property is present in the beginning of the recovery phase, transmitting these packets wanted by a large number of receivers will remove a large number of vertices from the graph, increase the number of edges, and thus increase the coding density. At the same time, the number of receivers, still wanting these packets, will considerably decrease after each transmission. Consequently, the MoWPS strategy will result in serving a smaller number of vertices in subsequent transmissions, which may decrease the coding density at some stages of the recovery transmission phase. We will illustrate and interpret these effects in Section VIII.

VI. Expected Coding Opportunity Evolution

Due to the great dependency of the coding opportunity evolution expression in (3) on whether the selected packets are widely wanted or not, it is difficult to infer from it any receiver selection strategies that densify the coding opportunities. Nonetheless, finding efficient receiver selection strategies is crucial in optimizing several IDNC parameters, such as the completion delay [13] and decoding delay [14]. Thus, it is very important to find another representation of the edge set size evolution that can be used to identify the best receiver selection strategy to densify the number of coding opportunities.

To reach this goal, we will employ the expected value representation of the edge set size, which only depends on the cardinalities of the receivers' feedback sets and eliminates the dependency on their actual packet contents. In this case, each Wants set $\mathcal{W}_i$ becomes a random set of packets of size ψ_i drawn from the pool of N original source packets. Consequently, the number of packets that are found in both Wants sets of two receivers i and k (i.e. $|\mathcal{W}_i \cap \mathcal{W}_k|$) becomes a random variable with hypergeometric

distribution. With this approach, the expression of the edge set size evolution will be an expectation given these hypergeometric random variables. We will derive this expression in the following two theorems.

Theorem 4. *Given the receivers' feedback set cardinalities, the expected edge set cardinality of the graph is equal to:*

$$\mathbb{E}\left[|\mathcal{E}|\right] = \frac{1}{2}\sum_{i=1}^{M}\psi_i\mathbb{E}\left[\Delta_i\right] = \frac{1}{2}\sum_{i=1}^{M}\psi_i\left\{\sum_{\substack{k=1\\k\neq i}}^{M}\frac{\psi_k}{N}\left(1+\frac{\varrho_k\varrho_i}{N-1}\right)\right\}\,, \tag{5}$$

where $\mathbb{E}\left[\Delta_i\right]$ *is the expected degree of vertices induced by receiver* i*.*

Proof: The proof can be found in Appendix D. ■

Theorem 5. *For any given feedback state and any given maximal clique* κ*, chosen for transmission at time* t*, the expected edge set size of the IDNC graph at time* $t+1$ *is expressed as:*

$$\mathbb{E}\left[\left|\mathcal{E}^{(t+1)}\right|\right] = \mathbb{E}\left[\left|\mathcal{E}^{(t)}\right|\right] - \frac{1}{2}\sum_{i\in\mathcal{T}}q_i\mathbb{E}\left[\Delta_i^{(t)}\right] + \frac{1}{2}\sum_{i\in\mathcal{T}}\psi_i\left(\alpha_i - \frac{q_i\gamma_i}{\psi_i}\right) + \frac{1}{2}\sum_{i\notin\mathcal{T}}\psi_i\beta_i\,, \tag{6}$$

where

$$\alpha_i = \sum_{\substack{k=1\\k\neq i}}^{M}q_i\xi_k - \sum_{\substack{k\in\mathcal{T}\\k\neq i}}\Phi_k(q_i) \qquad \beta_i = -\sum_{\substack{k\in\mathcal{T}\\k\neq i}}\Phi_k(0) \qquad \gamma_i = \sum_{\substack{k=1\\k\neq i}}^{M}\xi_k - \sum_{\substack{k\in\mathcal{T}\\k\neq i}}\Phi_k(1) \tag{7}$$

$$\Phi_k(x) = \frac{q_k}{N}\left(1+\frac{\left(\varrho_k-\psi_k+1\right)\left(\varrho_i+x\right)}{N-1}\right) \qquad \xi_k = \frac{\psi_k\varrho_k}{N(N-1)} \tag{8}$$

Proof: The proof can be found in Appendix E. ■

The above formula in (6) shows that the expected edge set size at $t+1$ is affected by two main components with respect to its value at time t. First, it suffers from a reduction due to the potential disappearance of served vertices and their edges, which is quantified by $\frac{1}{2}q_i\mathbb{E}\left[\Delta_i^{(t)}\right]$ for each targeted receiver. The second component is the change in expected degrees of the remaining vertices, which is quantified by $\frac{1}{2}\psi_i\left(\alpha_i - \frac{q_i\gamma_i}{\psi_i}\right)$ or $\frac{1}{2}\psi_i\beta_i$ for each targeted or non-targeted receiver, respectively. We can thus investigate the receiver selection strategy, maximizing the expected edge set size evolution, by studying these two components in the next section.

VII. Receiver Selection Strategies

To investigate the receiver selection strategy, maximizing the number of coding opportunities, we will analyze two main components affecting it.

A. Vertex Disappearance

The disappearance of vertices and its attributed loss of their adjacent edges is a natural outcome of targeting the packet requests of the different receivers throughout the recovery transmission process and is unavoidable. Nonetheless, we can still reduce the effect of this loss component by serving the vertices with smaller degrees. The following theorem compares the expected vertex degrees of two receivers given the sizes of their Wants sets.

Theorem 6. *If $\psi_i > \psi_h$, then $\mathbb{E}[\Delta_h] > \mathbb{E}[\Delta_i]$.*

Proof: The proof can be found in Appendix F. ■

Now, if $q_i < q_h$ and $\psi_i > \psi_h$, $\frac{1}{2}q_i\mathbb{E}[\Delta_i] < \frac{1}{2}q_h\mathbb{E}[\Delta_h]$. Consequently, serving receivers with largest Wants sets and erasure probabilities results in a smaller loss in the resulting edge set size.

B. Degrees of Remaining Vertices

To study the factors affecting the evolution of the degrees of remaining vertices, we introduce the following theorem.

Theorem 7. *If $\psi_i > 0$, $\psi_k > 1$, $\forall\ k \in \mathcal{T} \setminus i$ and $\psi_k \leq \varrho_k$, $\forall\ k \in \mathcal{T} \setminus i$, then $\alpha_i - \frac{q_i\gamma_i}{\psi_i} \geq \beta_i$ for any $i \in \mathcal{M}$.*

Proof: The proof can be found in Appendix 7. ■

The above theorem implies that the increase in the degrees of the remaining vertices of any receiver is larger when it is targeted than when it is not. Thus, moving a receiver i from the non-targeted set to the targeted set results in adding $\frac{\psi_i}{2}\left(\alpha_i - \frac{q_i\gamma_i}{\psi_i} - \beta_i\right) \geq 0$ edges to $\mathcal{G}$. This term is larger when ψ_i is larger, and thus moving a receiver with a larger Wants set to the targeted receiver set adds more edges to the primary graph than moving a receiver with a smaller Wants set. Consequently, a larger increase in the expected edge set size is obtained when targeting the maximum number of receivers having larger Wants sets.

Another important insight about the values of α_i and β_i can be inferred from a closer look at their components $\Phi_k(q_i)$ and $\Phi_k(0)$, respectively. Since the terms $\sum_{\substack{k\in\mathcal{T}_\rho(\kappa)\\k\neq i}}\Phi_k(q_i)$ and $\sum_{\substack{k\in\mathcal{T}_\rho(\kappa)\\k\neq i}}\Phi_k(0)$ are subtractive terms from α_i and β_i, respectively, selecting the receivers, with smaller $\Phi_k(q_i)$ and $\Phi_k(0)$, as targeted receivers, increases the values of α_i and β_i, respectively. Now, if $q_k < q_h$, $\psi_k > \psi_h$, we have:

$$q_k\left(\varrho_k - \psi_k + 1\right) < q_h\left(\varrho_h - \psi_h + 1\right) \qquad \Rightarrow \qquad \Phi_k(q_i) < \Phi_h(q_i) \quad \text{and} \quad \Phi_k(0) < \Phi_h(0)\,. \tag{9}$$

Consequently, the receivers having larger Wants sets and erasure probabilities have smaller values of $\Phi_k(q_i)$ and $\Phi_k(0)$, and thus targeting them increases α_i and β_i. This result also makes the term $\frac{q_i\gamma_i}{\psi_i}$ negligible in the $\left(\alpha_i - \frac{q_i\gamma_i}{\psi_i}\right)$ term for targeted receivers having larger Wants sets and erasure probabilities.

From Theorem 7 and the above observations on the values of α_i's and β_i's, we can conclude that targeting the maximum number of receivers with largest Wants sets and erasure probabilities results in a larger increase in the degrees of the remaining vertices in the IDNC graph.

C. Overall Strategy

From the above two sections, we can conclude that both factors identified from Theorem 2 achieve a larger contribution in the number of edge set size at time $t+1$, with respect to its value at time t, when the sender targets the maximum number of receivers with largest Wants sets and erasure probabilities. We will refer to this strategy as the *Worst Receiver Targeting (WoRT) strategy*. We can express this strategy as choosing the maximal clique κ^* in each transmission such that:

$$\kappa^* = \arg\min_{\kappa\in\mathcal{G}} \sum_{i|v_{ij}\in\kappa} \left(\frac{\psi_i}{q_i}\right)^n , \tag{10}$$

where n is a biasing factor.

VIII. Simulations

In this section, we test, through simulations, the performances of our identified strategies in maximizing the coding density of the system during the transmission of a frame, and compare them to other well-known strategies. The simulation scenario consists of M receivers having different packet erasure probabilities while maintaining the average erasure probability (ε) constant. These erasure probabilities are assumed to be fixed during the transmission of a frame but change from frame to frame during the simulation. The tested strategies in this simulations are:

- RND: Random clique selection [3].
- MC: Maximum clique selection [16].
- MWC-R: Maximum weighted clique (MWC) selection, in which the weight of vertex v_{ij} is defined as the reception probability of receiver i.
- MoWPS: Maximum weight clique selection defined as in (4).
- WoRT: Maximum weight clique selection defined as in (10).

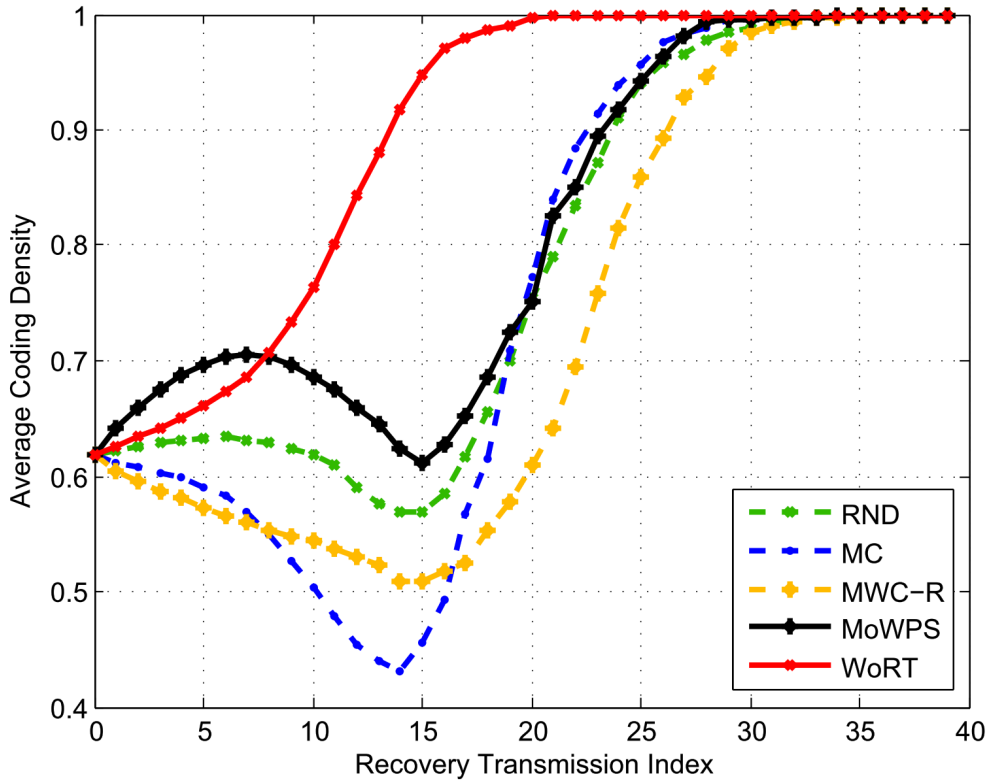

Fig. 3. Average coding density evolution for $M = 50$ and $N = 20$

All figures represent the average coding density of the graph after each transmission, when the corresponding strategy is employed during all the recovery transmission phase. This average is computed over a large number of iterations, in each of which we compute the graph density after each transmission. We then average all the densities evaluated at the same transmission index.

A. Results with Optimal MWC Algorithm

Figure 3 depicts the coding density evolution inside the IDNC graph for $M = 50$ and $N = 20$. The erasure probabilities of different receivers take values in the range of 0.01 to 0.3, such that $\varepsilon = 0.15$. From Figure 3, we can draw the following observations. As expected, the MoWPS strategy considerably increases the coding density for the first 25% of the recovery transmissions (i.e. transmissions 1 to 8) due to the presence of packets requested by a large number of receivers during this period. However, due to the reduction of these numbers after several transmissions, the strategy cannot continue increasing the coding opportunities. Indeed, the restriction of serving such packets when there are none forces the algorithm to serve less vertices, which results in a smaller coding density, as shown in the intermediate 22% of the recovery phase (i.e. transmissions 9 to 15). Towards the end of the phase, the number of

requests is naturally reduced and the Has sets become very close to N. This naturally increases the coding opportunities due to Condition C2 in Section III. Consequently, the remaining transmissions target most of the receivers, which both further increases the coding opportunities and reduces the number of vertices in the graph, thus increasing the coding density.

As for the effect of receiver selection, we can clearly see that the WoRT strategy considerably outperforms all other receiver selection strategies. Moreover, we can see that it monotonically increases the coding density, which proves its ability to increase the coding density over the full course of the frame transmission. This performance is supported by the fact that a single transmission can reduce the Wants set of any receiver by at most one. Consequently, there will always exist some receivers with larger Wants set sizes during most of the recovery phase, which can be targeted to continuously increase the number of coding opportunities and the coding density.

Another important result is that the MC strategy, serving the maximum number of receivers (or requests) in each transmission, and widely considered in most opportunistic network coding works, results in a very bad evolution of coding density. We can infer from Figure 3 that this strategy serves very few large cliques in the very beginning of the recovery phase and then is left with smaller cliques to serve, until it is close to the completion of frame delivery. This can be explained from (31), showing the high adjacency of vertices with smaller Wants sets. Consequently, the MC strategy mostly targets receivers with smaller Wants sets, which clearly reduces the edge set size according to Section VII. Despite the common intuition that this strategy can optimize different network coding parameters, this result shows that it may not be able to truly do so. We will show one example of this fact in the next section.

Figures 4 and 5 depict the coding density evolution inside the IDNC graph for different values of M (when $N = 20$ and $\varepsilon = 0.15$) and N (when $M = 50$ and $\varepsilon = 0.15$), respectively.

We can see from both figures that all the observation and conclusions deduced from Figure 3 hold for various values of M and N.

B. Results with Greedy MWC Algorithm

Since the optimal solution of MWC selection problem is NP-hard, we test the performance of a greedy MWC algorithm, which adds the vertex with the highest weight to the output clique in each iteration. For a better representation of the adjacency effect on the vertex selection in this greedy approach, we modify

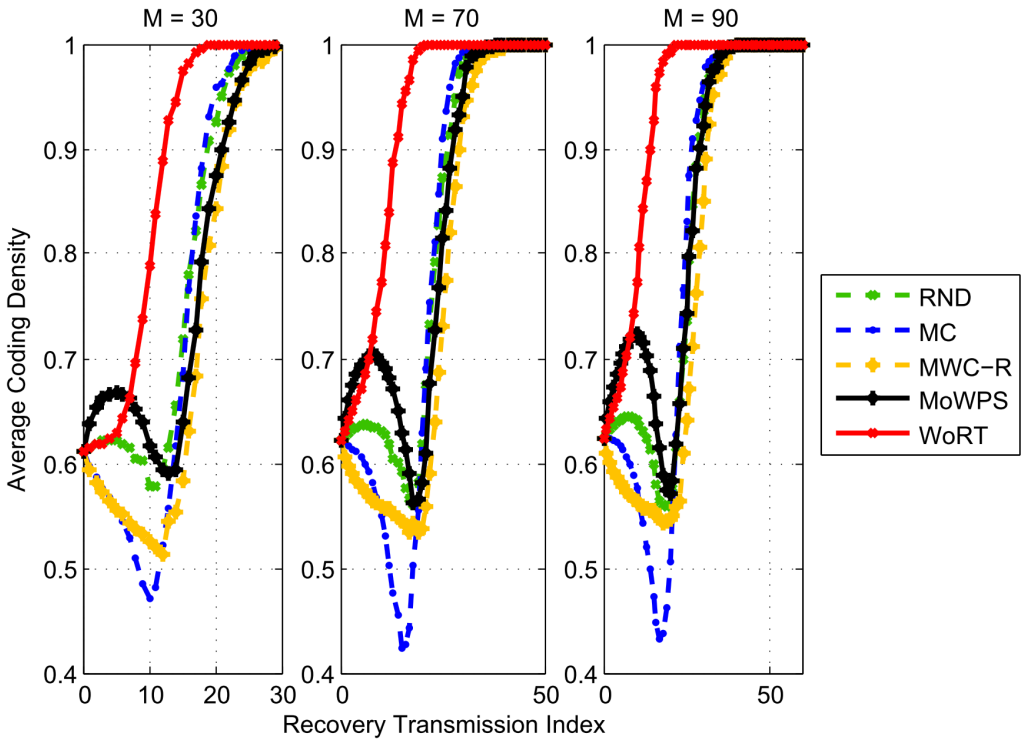

Fig. 4. Average coding density evolution for different values of M, when $N = 20$ and $\varepsilon = 0.15$

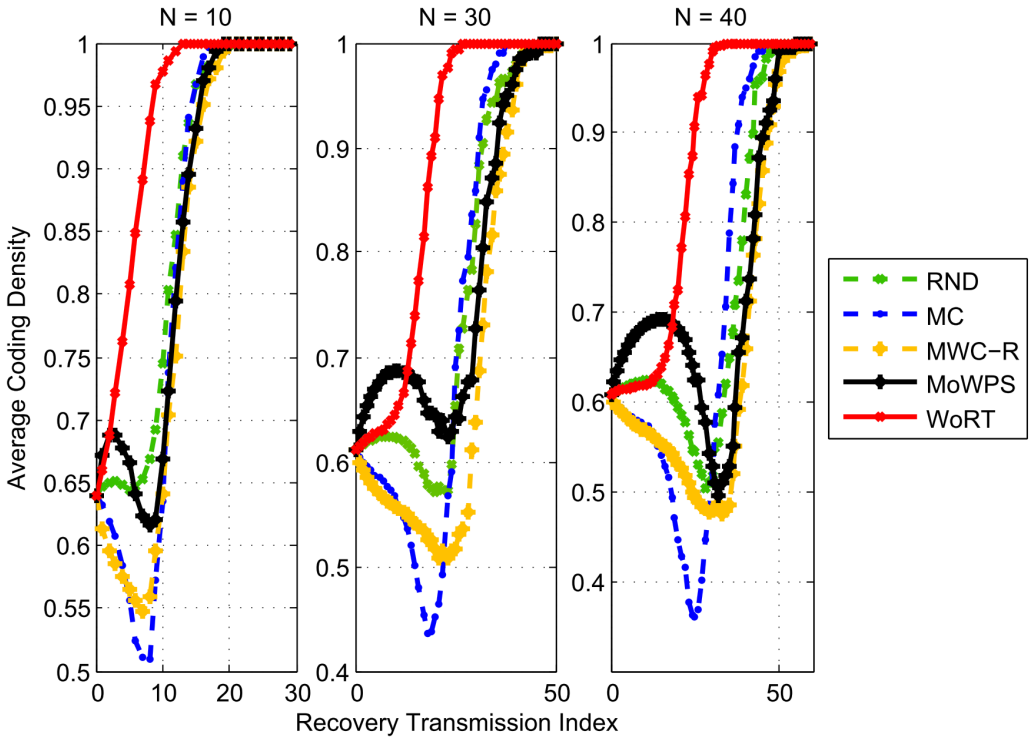

Fig. 5. Average coding density evolution for different values of N, when $M = 50$ and $\varepsilon = 0.15$

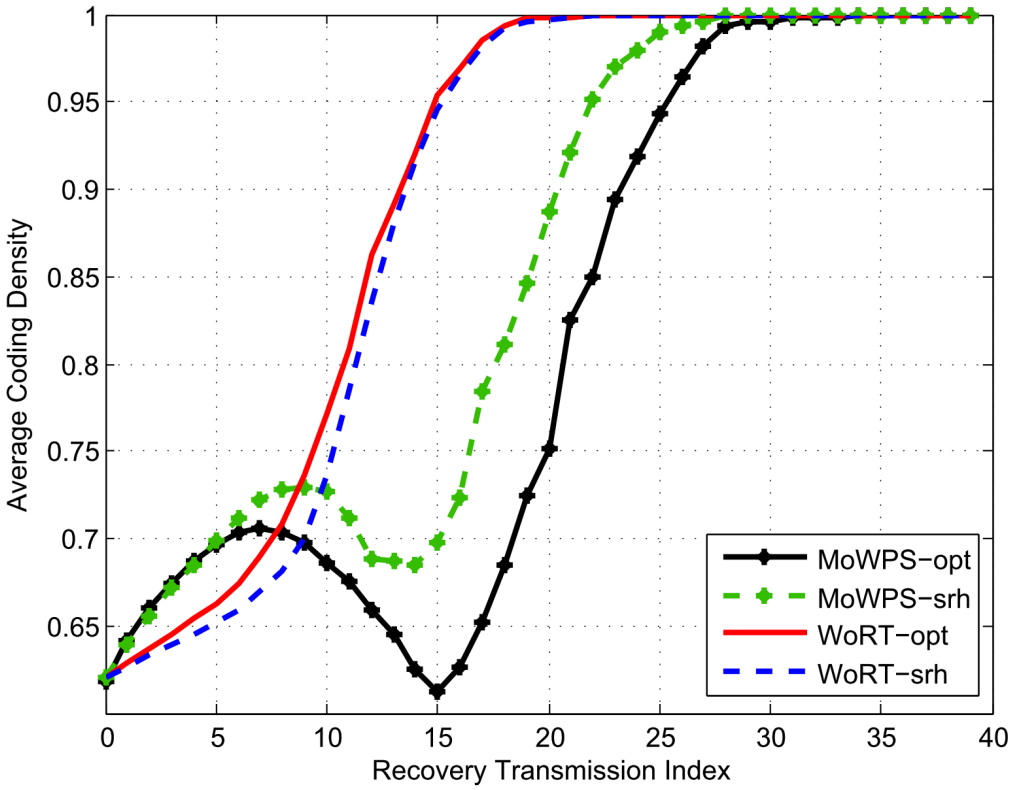

Fig. 6. Average coding density evolution for $M = 50$ and $N = 20$.

the weight of each vertex v_{ij}, having original weight w_{ij}, to be:

$$\omega_{ij} = w_{ij} \cdot \sum_{\forall\ v_{kl} \in \mathcal{V}} I_{v_{kl} \in \mathcal{V}_{ij}} \cdot w_{kl} \tag{11}$$

where $\mathcal{V}_{ij}$ is the set of adjacent vertices to v_{ij} and where I_x is an indicator function, which is equal to one if x is true and zero otherwise. Consequently, these new vertex weights reflect not only their individual weights w_{ij} but also their adjacency to a large number of vertices having high individual weights. Since these modified weights need to be recomputed after each iteration to reflect the new adjacency conditions, the complexity of this greedy algorithms is $O(M^2N)$.

Figure 6 depicts the coding density evolution of optimal and greedy MWC algorithms for $M = 50$, $N = 20$ and $\varepsilon = 0.15$. In this figure, we only consider the MoWPS and WoRT strategies. For the WoRT strategy, we can see that the greedy algorithm achieves the same monotonically increasing performance as the optimal algorithm, with a slight gap between them. For the MoWPS strategy, we notice that the greedy algorithm follows the same trend of the optimal algorithms but achieves larger coding density after the first 7 transmissions. This can be explained by the fact that the greedy algorithm loses some chances of serving packets with largest demands in the first few transmissions, compared to the optimal algorithm.

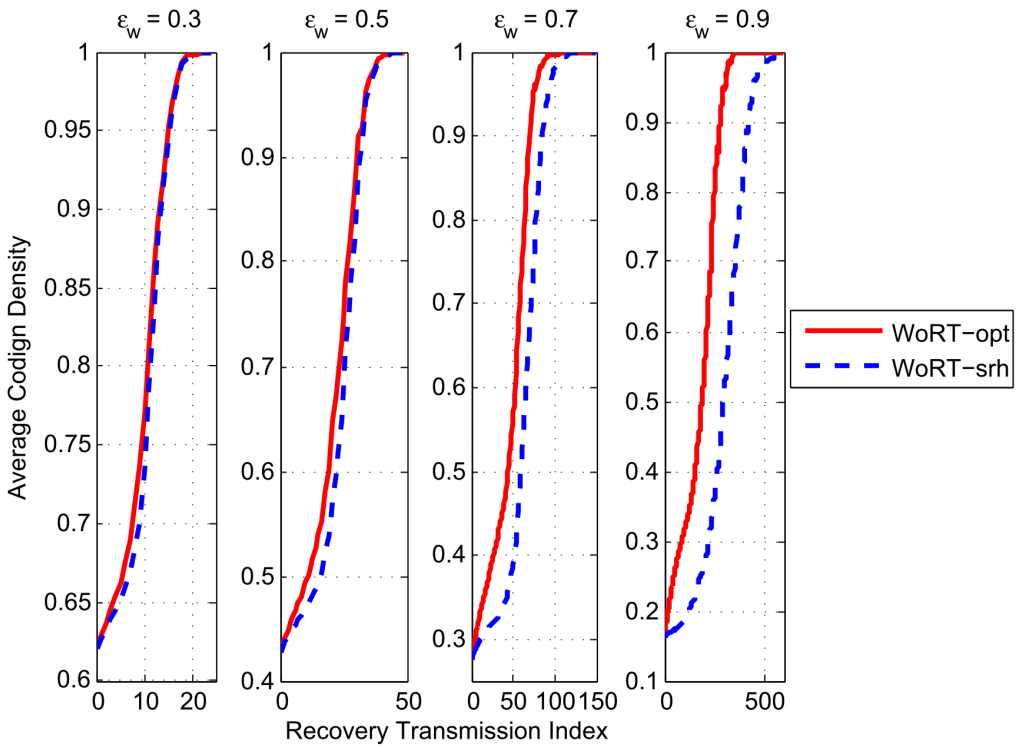

Fig. 7. Average coding density evolution for different ε_w ($M = 50$ and $N = 20$).

This luckily makes it less affected by the degradation phenomenon that happens to the optimal algorithm, as explained above. This effect both preserves better coding opportunities for the greedy algorithm in the intermediate range of transmissions and allows a faster boost up towards the end.

C. *Effect of Erasure Probabilities*

Since the receiver selection in the WoRT strategy greatly depends on their erasure probabilities, we test its performance for several erasure probabilities $\varepsilon_w = [0.3, 0.5, 0.7, 0.9]$ in Figure 7. We can see from the figure that the performances of both the optimal and greedy MWC algorithms always achieve the same monotonically increasing trend of average coding densities for all these diverse values of erasure probabilities. We also notice a larger delay between the optimal and greedy algorithms as the erasure probability increases. This result is expected as the greedy algorithm will naturally degrade more in the overall performance as it runs for a larger number of recovery transmissions at high erasure probabilities.

IX. Case Study: Completion Delay

In this section, we study the effect of maximizing the coding density on reducing the average completion delay (i.e. the number of recovery transmissions) in IDNC. Intuitively, one can think that the optimal

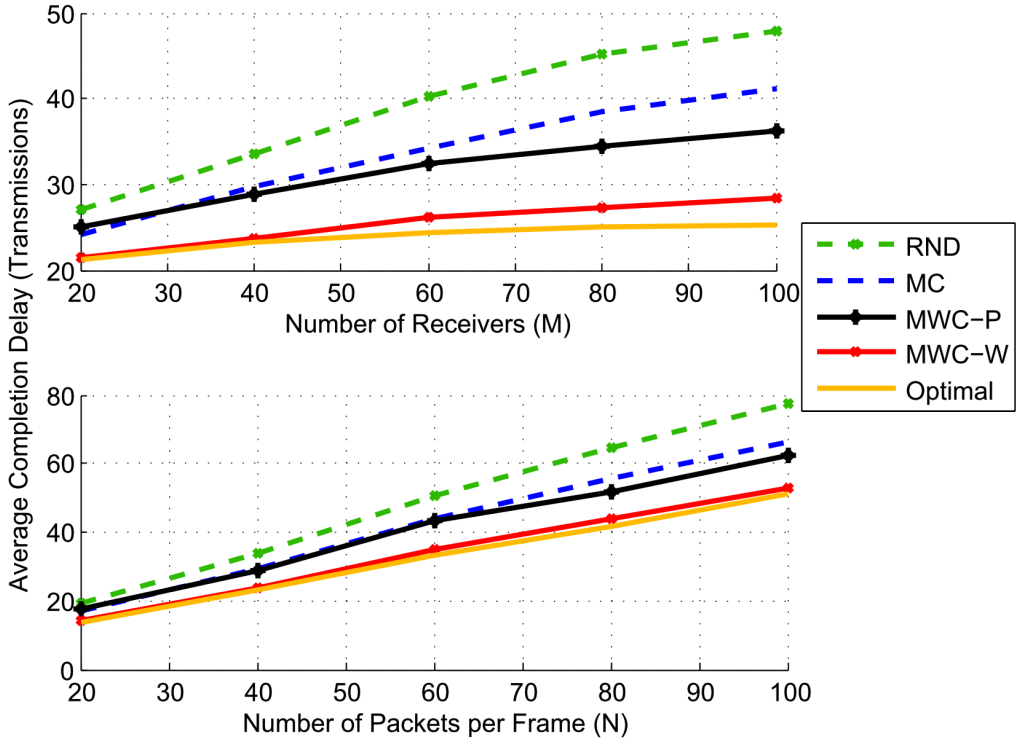

Fig. 8. Comparison of average completion delays against M and N.

IDNC completion delay can be achieved by serving the maximum number of vertices in each recovery transmission (i.e. the MC strategy), as this should apparently deplete the graph faster. However, we have shown that the MC strategy suffers from severe degradation in the coding density compared to the WoRT strategy. We will thus compare the completion delay of these two strategies to see whether increasing the coding density in earlier transmissions is important in reducing the frame completion delay.

Figure 8 depicts the comparison of the average completion delays achieved by the RND, MC, MoWPS and WoRT algorithms to global optimal completion delay over all linear network codes. The upper subfigure illustrates this comparison against M, for $N = 40$ and $\varepsilon = 0.15$, whereas the lower one compares the performances against N for $M = 40$ and $\varepsilon = 0.15$. Both subfigures show that the WoRT strategy considerably outperforms all other strategies including the MC strategy, especially as M and N increase. We can also see that WoRT strategy achieves a near-optimal completion delay performance. These results clearly show that the MC strategy cannot achieve a low completion delay, due to the effect explained in Section VIII. On the other hand, the WoRT strategy, which serves relatively smaller cliques in the beginning of the recovery phase, achieves a better completion delay as it persistently increases the coding density in the graph. Thus, it always finds large cliques to serve in later recovery transmissions, and thus complete their delivery much faster.

X. Conclusion

In this paper, we investigated the receiver and packet selection strategies that densify the IDNC coding opportunities in wireless broadcast. We first derived an expression for the exact evolution of the edge set size in the IDNC graph, after the transmission of any arbitrary coded packet. From this expression, we showed that serving packets requested by a larger number of receivers tends to maximize the coding density, especially in the beginning of the recovery transmission phase. We then derived an expected expression of the edge set size evolution, after ignoring the feedback set contents and keeping their cardinalities. We employed this expression to show that targeting the maximum number of receivers, having the largest Wants sets and erasure probabilities, tends to both maximize the expected number of coding opportunities and monotonically increase the expected coding density. Simulation results justified our theoretical findings. Finally, we demonstrated the validity and importance of increasing the coding density through a case study on the IDNC completion delay. The study showed that the identified WoRT strategy achieves a significantly lower completion delay than the MC strategy (which was intuitively expected to perform better), due to its inability to preserve high coding density throughout the frame transmission. We thus recommend the observance of the coding density increasing strategies when other network parameters are optimized over the full course of a frame transmission.

Appendix A
Proof of Theorem 1

It is well known from graph theory that the edge set size of any graph is equal to half the sum of its vertex degrees. Consequently, we will first derive the expression for the degree of an arbitrary vertex v_{ij} in the graph. From the adjacency conditions C1 and C2 in Section III, we conclude the following facts:

- Vertex v_{ij} is not adjacent to any vertex of the same receive i.
- If $j \in \mathcal{W}_k$, v_{ij} cannot be adjacent to any vertex of receiver k due to violation of C2, except for vertex v_{kj} which arises from C1.
- If $j \in \mathcal{H}_k$, v_{ij} can be adjacent to any vertex of receiver k (induced from $\mathcal{W}_k$), except for all vertices v_{kl} for which $l \notin \mathcal{H}_i \Rightarrow l \in \mathcal{W}_i$.

From these facts, we can express the degree of a vertex v_{ij} as follows:

$$\Delta_{ij} = \sum_{\substack{k=1 \\ k \neq i}}^{M} I_{j \in \mathcal{W}_k} + I_{j \in \mathcal{H}_k} \left(|\mathcal{W}_k| - |\mathcal{W}_k \cap \mathcal{W}_i| \right) \tag{12}$$

TABLE I
EVOLUTION CASES OF THE WANTS SETS OF A PAIR OF RECEIVERS i AND k WHEN THE SENDER TRANSMITS AN ARBITRARY CLIQUE κ TARGETING A SET OF RECEIVERS $\mathcal{T}$

Cases			ψ_i	ψ_k	ψ_{ik}
i	k	p_i/p_k			
$\notin \mathcal{T}$	$\notin \mathcal{T}$	-	ψ_i	ψ_k	ψ_{ik}
$\notin \mathcal{T}$	$\in \mathcal{T}$	$p_k \notin \mathcal{W}_i$	ψ_i	$\psi_k - X_k$	ψ_{ik}
$\notin \mathcal{T}$	$\in \mathcal{T}$	$p_k \in \mathcal{W}_i$	ψ_i	$\psi_k - X_k$	$\psi_{ik} - X_k$
$\in \mathcal{T}$	$\notin \mathcal{T}$	$p_i \notin \mathcal{W}_k$	$\psi_i - X_i$	ψ_k	ψ_{ik}
$\in \mathcal{T}$	$\notin \mathcal{T}$	$p_i \in \mathcal{W}_k$	$\psi_i - X_i$	ψ_k	$\psi_{ik} - X_i$
$\in \mathcal{T}$	$\in \mathcal{T}$	$p_i \notin \mathcal{W}_k$	$\psi_i - X_i$	$\psi_k - X_k$	ψ_{ik}
$\in \mathcal{T}$	$\in \mathcal{T}$	$p_i \in \mathcal{W}_k$	$\psi_i - X_i$	$\psi_k - X_k$	$\psi_{ik} - X_{ik}$

Now, the sum $\Sigma\Delta_i$ of all the degrees of the vertices induced by receiver i can be expressed as:

$$\Sigma\Delta_i = \sum_{j\in\mathcal{W}_i}\sum_{\substack{k=1\\k\neq i}}^{M} I_{j\in\mathcal{W}_k} + I_{j\in\mathcal{H}_k}\left(|\mathcal{W}_k| - |\mathcal{W}_k\cap\mathcal{W}_i|\right) = \sum_{\substack{k=1\\k\neq i}}^{M}\left(\sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{W}_k}\right) + \sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{H}_k}\cdot\left(\psi_k - |\mathcal{W}_k\cap\mathcal{W}_i|\right)$$
$$= \sum_{\substack{k=1\\k\neq i}}^{M}|\mathcal{W}_i\cap\mathcal{W}_k| + \sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{H}_k}\cdot\left(\psi_k - |\mathcal{W}_k\cap\mathcal{W}_i|\right)\ . \tag{13}$$

The expression $\sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{H}_k}$ can be easily shown to be equal to $|\mathcal{W}_i\cap\mathcal{H}_k|$. We can also easily infer that $\mathcal{W}_i\cap\mathcal{H}_k = \mathcal{W}_i\setminus(\mathcal{W}_i\cap\mathcal{W}_k)$. Consequently, we get:

$$\sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{H}_k} = |\mathcal{W}_i\cap\mathcal{H}_i| = |\mathcal{W}_i| - |\mathcal{W}_i\cap\mathcal{W}_k|\ . \tag{14}$$

From (14), (13) and using the definitions in (2), we get:

$$|\mathcal{E}| = \frac{1}{2}\sum_{i=1}^{M}\Sigma\Delta_i = \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M}\left[\psi_{ik} + (\psi_i - \psi_{ik})(\psi_k - \psi_{ik})\right] = \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M}\left(\psi_{ik} + \theta_{ik}\theta_{ki}\right)\ . \tag{15}$$

APPENDIX B
PROOF OF THEOREM 2

When an arbitrary clique κ, with a given set of targeted receivers $\mathcal{T}$, is chosen for transmission, the values of ψ_i, ψ_k and ψ_{ik} in (2) change according to the cases depicted in Table I for each pair of receivers. In this table, X_{ik} is a joint indicator function, which is equal to 1 if either i or k received the transmitted packet and is zero otherwise. It is easy to see that X_{ik} can be expressed as $X_{ik} = X_i + X_k - X_iX_k$. Let $Y_{ik}^{(t)} = \psi_{ik} + \theta_{ik}\theta_{ki}$. According to the possible changes in Table I, the double summation in (2), can be divided into these seven categories. In each category, the values of ψ_i, ψ_k and ψ_{ik} in (2) are replaced

in their $Y_{ik}^{(t)}$ expression by their corresponding evolved values, depicted in the table, to reach to $Y_{ik}^{(t+1)}$. Applying these changes, using the definition $X_{ik} = X_i + X_k - X_iX_k$, and re-arranging the terms, we get the following results for each category.

1) For $i \notin \mathcal{T}$, $k \notin \mathcal{T}$, there will be no change in $Y_{ik}^{(t)}$.
2) For $i \notin \mathcal{T}$, $k \in \mathcal{T}$ and $p_k \notin \mathcal{W}_i$, $Y_{ik}^{(t+1)} = \psi_{ik} + \theta_{ik}\left(\theta_{ki} - X_k\right) = Y_{ik}^{(t)} - X_k\theta_{ik}$.
3) For $i \notin \mathcal{T}$, $k \in \mathcal{T}$ and $p_k \in \mathcal{W}_i$, $Y_{ik}^{(t+1)} = \psi_{ik} - X_k + \left(\theta_{ik} + X_k\right)\theta_{ki} = Y_{ik}^{(t)} + X_k\hat{\theta}_{ki}$.
4) For $i \in \mathcal{T}$, $k \notin \mathcal{T}$, we get the same results in the points 2 and 3, for $p_i \notin \mathcal{W}_k$ and $p_i \in \mathcal{W}_k$, respectively, by replacing index k by i and vice versa in the terms X_k, θ_{ik} and $\hat{\theta}_{ki}$.
5) For $i \in \mathcal{T}$, $k \in \mathcal{T}$ and $p_i \notin \mathcal{W}_k$, $Y_{ik}^{(t+1)} = \psi_{ik} + \left(\theta_{ik} - X_i\right)\left(\theta_{ki} - X_k\right) = Y_{ik}^{(t)} + X_iX_k - X_i\theta_{ki} - X_k\theta_{ik}$.
6) For $i \in \mathcal{T}$, $k \in \mathcal{T}$ and $p_i \in \mathcal{W}_k$,

$$\begin{aligned} Y_{ik}^{(t+1)} &= \psi_{ik} - X_i - X_k + X_iX_k + \left(\theta_{ik} + X_k - X_iX_k\right)\left(\theta_{ki} + X_i - X_iX_k\right) \\ &= Y_{ik}^{(t)} + X_i\hat{\theta}_{ik} + X_k\hat{\theta}_{ki} - X_iX_k\left(\hat{\theta}_{ik} + \hat{\theta}_{ki} - X_iX_k\right) . \end{aligned}$$

The theorem follows by substituting the above equations in (2).

APPENDIX C
PROOF OF THEOREM 3

According to the targeting status of any two receivers i and k, they will have a pairwise entry in only one of the summations in (3). If $i \notin \mathcal{T}$ and $k \in \mathcal{T}$, the number of pairwise edges at $t+1$ $\left(Y_{ik}^{(t+1)}\right)$ is increased by $X_k\hat{\theta}_{ki}$ when $p_k \in \mathcal{W}_i$ and is reduced by $X_k\theta_{ik}$ when $p_k \notin \mathcal{W}_i$. If $i \in \mathcal{T}$ and $k \notin \mathcal{T}$, $Y_{ik}^{(t+1)}$ is increased by $X_i\hat{\theta}_{ik}$ when $p_i \in \mathcal{W}_k$ and is reduced by $X_i\theta_{ki}$ when $p_i \notin \mathcal{W}_k$.

If both receivers are targeted, we have one of the following cases:

Case 1: $X_i = 0$ and $X_k = 0$: $Y_{ik}^{(t+1)} = Y_{ik}^{(t)}$ whether $p_i \in \mathcal{W}_k$ or not.

Case 2: $X_i = 1$ and $X_k = 0$:

1) For $p_i \notin \mathcal{W}_k$, $Y_{ik}^{(t+1)} = Y_{ik}^{(t)} - \theta_{ki} \leq Y_{ik}^{(t)} - 1$.
2) For $p_i \in \mathcal{W}_k$, $Y_{ik}^{(t+1)} = Y_{ik}^{(t)} + \hat{\theta}_{ik}$.

Case 3: $X_i = 0$ and $X_k = 1$:

1) For $p_i \notin \mathcal{W}_k$, $Y_{ik}^{(t+1)} = Y_{ik}^{(t)} - \theta_{ik} \leq Y_{ik}^{(t)} - 1$.
2) For $p_i \in \mathcal{W}_k$, $Y_{ik}^{(t+1)} = Y_{ik}^{(t)} + \hat{\theta}_{ki}$.

Case 4: $X_i = 1$ and $X_k = 1$:

1) For $p_i \notin \mathcal{W}_k$, $Y_{ik}^{(t+1)} = Y_{ik}^{(t)} - \theta_{ki} - \theta_{ik} + 1 \leq Y_{ik}(t) - 1$.

2) For $p_i \in \mathcal{W}_k$, $Y_{ik}^{(t+1)} = Y_{ik}^{(t)} - 1$.

The two inequalities in Condition 1 of both Cases 2 and 3 (i.e. when $p_i \notin \mathcal{W}_k$) arise from the following two facts:

1) p_i is in $\mathcal{W}_i$ but not $\mathcal{W}_k$.

2) Both receivers are targeted.

Fact 1 implies that there must exists at least one packet satisfying this condition and thus:

$$\theta_{ik} = \psi_i - \psi_{ik} = |\mathcal{W}_i \setminus \mathcal{W}_k| \geq 1. \tag{16}$$

Fact 2 implies that k is targeted by packet p_k such that:

- $p_k \neq p_i$: or else p_i must be in $\mathcal{W}_k$, which contradicts with Condition 1 of both Cases 2 and 3.
- $p_k \notin \mathcal{W}_i$: or else the combination $p_i \oplus p_k$ will not be instantly decodable at i, which contradicts with the fact that i is targeted.

Consequently, there exists at least one packet p_k which is in $\mathcal{W}_k$ but not $\mathcal{W}_i$ and thus

$$\theta_{ki} = \psi_k - \psi_{ik} = |\mathcal{W}_k \setminus \mathcal{W}_i| \geq 1. \tag{17}$$

The last inequality in Condition 1 of Case 4 for $p_i \notin \mathcal{W}_k$ arises from the fact that packet p_i is in $\mathcal{W}_i$ but not in $\mathcal{W}_k$ and thus cannot be part of their intersection. Consequently, $\psi_i \geq \psi_{ik} + 1 \;\Rightarrow\; \theta_{ik} \geq 1$ and $\psi_k \geq \psi_{ik} + 1 \;\Rightarrow\; \theta_{ki} \geq 1$ and thus the left hand term will always be less than or equal to $Y_{ik}^{(t)} - 1$.

APPENDIX D
PROOF OF THEOREM 4

To prove this theorem, we need to introduce this lemma, which is proved in Appendix H.

Lemma 1. *For any given feedback state, the expected degree of any of the vertices induced by receiver i (denoted by Δ_i) is equal to:*

$$\mathbb{E}\left[\Delta_i\right] = \sum_{\substack{k=1 \\ k \neq i}}^{M} \frac{\psi_k}{N} \left(1 + \frac{\varrho_k \varrho_i}{N-1}\right) . \tag{18}$$

We will start our proof using (13). When we ignore the contents of the different sets, we can derive

an expression for the expected edge set size of the graph as follows:

$$\mathbb{E}\left[|\mathcal{E}|\right] = \frac{1}{2}\sum_{i=1}^{M}\mathbb{E}\left[\Sigma\Delta_i\right] = \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M}\mathbb{E}\left[|\mathcal{W}_i\cap\mathcal{W}_k|\right] + \psi_k\mathbb{E}\left[\sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{H}_k}\right] - \mathbb{E}\left[\sum_{j\in\mathcal{W}_i} I_{j\in\mathcal{H}_k}\cdot|\mathcal{W}_k\cap\mathcal{W}_i|\right]$$
$$= \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M}\frac{\psi_i\psi_k}{N} + \frac{\psi_k\psi_i\varrho_k}{N} - \sum_{j\in\mathcal{W}_i}\mathbb{E}\left[I_{j\in\mathcal{H}_k}\cdot|\mathcal{W}_k\cap\mathcal{W}_i|\right] \tag{19}$$

Note that the indicator function in the last term can be only zero or one. Consequently, the expectation of its multiplication with $|\mathcal{W}_k\cap\mathcal{W}_i|$ can be only evaluated for $I_{j\in\mathcal{H}_k}=1$. In this case, packet j cannot be in the intersection of $\mathcal{W}_k$ and $\mathcal{W}_i$. Consequently, this intersection is possible only with the other ψ_i-1 packets of receiver i, and from the set of the remaining $N-1$ packets. Since the cardinality of the intersection of two sets of given sizes, whose elements are drawn of the same pool of $N-1$ elements, is a hypergeometric random variable, we have:

$$\mathbb{E}\left[I_{j\in\mathcal{H}_k}\cdot|\mathcal{W}_k\cap\mathcal{W}_i|\right] = \sum_{n=1}^{N-1} n\,\mathbb{P}\left[I_{j\in\mathcal{H}_k}=1, |\mathcal{W}_k\cap\mathcal{W}_i|=n\right]$$
$$= \sum_{n=1}^{N-1} n\,\mathbb{P}\left[|\mathcal{W}_k\cap\mathcal{W}_i|=n\Big|I_{j\in\mathcal{H}_k}=1\right]\cdot\mathbb{P}\left[I_{j\in\mathcal{H}_k}=1\right] = \sum_{n=1}^{N-1} n\frac{\binom{\psi_i-1}{n}\binom{N-1-\psi_i+1}{\psi_k-n}}{\binom{N-1}{\psi_k}}\,\frac{\varrho_k}{N} = \frac{\varrho_k\psi_k\left(\psi_i-1\right)}{N(N-1)}\,. \tag{20}$$

Substituting (20) in (19), we get:

$$\mathbb{E}\left[|\mathcal{E}|\right] = \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{k=1\\k\neq i}}^{M}\frac{\psi_i\psi_k}{N} + \frac{\psi_k\psi_i\varrho_k}{N} - \frac{\psi_i\varrho_k\psi_k(\psi_i-1)}{N(N-1)} = \frac{1}{2}\sum_{i=1}^{M}\psi_i\left\{\sum_{\substack{k=1\\k\neq i}}^{M}\frac{\psi_k}{N}\left[1+\varrho_k\left(1-\frac{\psi_i-1}{N-1}\right)\right]\right\}$$
$$= \frac{1}{2}\sum_{i=1}^{M}\psi_i\left\{\sum_{\substack{k=1\\k\neq i}}^{M}\frac{\psi_k}{N}\left(1+\frac{\varrho_k\varrho_i}{N-1}\right)\right\} = \frac{1}{2}\sum_{i=1}^{M}\psi_i\mathbb{E}\left[\Delta_i\right]\,. \tag{21}$$

APPENDIX E
PROOF OF THEOREM 5

To prove this theorem, we first need to introduce the following lemma, which is proved in Appendix I.

Lemma 2. *For a given maximal clique κ, chosen for transmission at time t, the expected degree of a*

receiver i vertex at time $t+1$ is expressed, for $i \in \mathcal{T}$ and $i \notin \mathcal{T}$, as:

$$\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\right] = \mathbb{E}\left[\Delta_i^{(t)}\right] + \sum_{\substack{k=1\\k\neq i}}^{M} q_i\xi_k - \sum_{\substack{k\in\mathcal{T}\\k\neq i}} \Phi_k(q_i) \tag{22}$$

$$\mathbb{E}\left[\Delta_{i\notin\mathcal{T}}^{(t+1)}\right] = \mathbb{E}\left[\Delta_i^{(t)}\right] - \sum_{k\in\mathcal{T}} \Phi_k(0) \tag{23}$$

When the maximal clique κ is chosen for transmission at time t, the receivers in $\mathcal{T}$ is targeted with the coded packet but are not guaranteed to receive that packet. From Lemma 4, we can derive the expression of the expected edge set size at time $t+1$, conditioned on the random vector $\mathbf{X}$ defined in Appendix I, as follows:

$$\mathbb{E}\left[\left|\mathcal{E}^{(t+1)}\right| \middle| \mathbf{X}\right] = \frac{1}{2}\sum_{i\in\mathcal{T}}(\psi_i - X_i)\,\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right] + \frac{1}{2}\sum_{i\notin\mathcal{T}}\psi_i\,\mathbb{E}\left[\Delta_{i\notin\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right] \ . \tag{24}$$

Now, taking the expectation operator over the random vector $\mathbf{X}$, we can get the expression for the expected edge set size at time $t+1$ as follows:

$$\begin{aligned}\mathbb{E}\left[\left|\mathcal{E}^{(t+1)}\right|\right] &= \mathbb{E}_{\mathbf{X}}\left\{\mathbb{E}\left[\left|\mathcal{E}^{(t+1)}\right| \middle| \mathbf{X}\right]\right\} \\ &= \frac{1}{2}\sum_{i\in\mathcal{T}}\psi_i\,\mathbb{E}_{\mathbf{X}}\left\{\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\} - \frac{1}{2}\sum_{i\in\mathcal{T}}\mathbb{E}_{\mathbf{X}}\left\{X_i\,\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\} + \frac{1}{2}\sum_{i\notin\mathcal{T}}\psi_i\,\mathbb{E}_{\mathbf{X}}\left\{\mathbb{E}\left[\Delta_{i\notin\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\}\end{aligned} \tag{25}$$

From (40), we know that:

$$\mathbb{E}_{\mathbf{X}}\left\{\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\} = \mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\right] = \mathbb{E}\left[\Delta_i^{(t)}\right] + \sum_{\substack{k=1\\k\neq i}}^{M} q_i\xi_k - \sum_{\substack{k\in\mathcal{T}\\k\neq i}}\Phi_k(q_i) = \mathbb{E}\left[\Delta_i^{(t)}\right] + \alpha_i \ . \tag{26}$$

Similarly, we know that:

$$\mathbb{E}_{\mathbf{X}}\left\{\mathbb{E}\left[\Delta_{i\notin\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\} = \mathbb{E}\left[\Delta_{i\notin\mathcal{T}}^{(t+1)}\right] = \mathbb{E}\left[\Delta_i^{(t)}\right] - \sum_{\substack{k\in\mathcal{T}\\k\neq i}}\Phi_k(0) = \mathbb{E}\left[\Delta_i^{(t)}\right] + \beta_i \ . \tag{27}$$

From (40), we can compute $\mathbb{E}_{\mathbf{X}}\left\{X_i\,\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\}$ as follows:

$$\begin{aligned}\mathbb{E}_{\mathbf{X}}\left\{X_i\,\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\} &= \mathbb{E}_{\mathbf{X}}\left\{X_i\right\}\mathbb{E}\left[\Delta_i^{(t)}\right] + \sum_{\substack{k=1\\k\neq i}}^{M}\frac{\psi_k\varrho_k\mathbb{E}_{\mathbf{X}}\left\{X_i^2\right\}}{N(N-1)} \\ &\quad - \sum_{\substack{k\in\mathcal{T}\\k\neq i}}\mathbb{E}_{\mathbf{X}}\left\{\frac{X_iX_k}{N} + \frac{\left(X_k\left(\varrho_k-\psi_k\right)+X_k^2\right)\left(\varrho_iX_i+X_i^2\right)}{N(N-1)}\right\}\end{aligned} \tag{28}$$

Using the definitions of ξ_k, $\Phi_k(x)$ and γ_i in (8) and (7), respectively, we get:

$$\mathbb{E}_{\mathbf{X}}\left\{X_i\ \mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\}=q_i\mathbb{E}\left[\Delta_i^{(t)}\right]+\sum_{\substack{k=1\\k\neq i}}^{M}q_i\xi_k-\sum_{\substack{k\in\mathcal{T}\\k\neq i}}q_i\Phi_k(1)=q_i\left(\mathbb{E}\left[\Delta_i^{(t)}\right]+\gamma_i\right)\ . \tag{29}$$

Finally, we know that:

$$\frac{1}{2}\sum_{i\in\mathcal{T}}\psi_i\mathbb{E}\left[\Delta_i^{(t)}\right]+\frac{1}{2}\sum_{i\notin\mathcal{T}}\psi_i\mathbb{E}\left[\Delta_i^{(t)}\right]=\mathbb{E}\left[\left|\mathcal{E}^{(t)}\right|\right] \tag{30}$$

The theorem follows by substituting (26), (27) and (28) in (25), re-arranging the terms and finally substituting (30) in the re-arranged equation.

APPENDIX F
PROOF OF THEOREM 6

Note that $\psi_i>\psi_h$ implicitly means that $\varrho_h<\varrho_i$. From Lemma 1 in Appendix D, we have:

$$\mathbb{E}\left[\Delta_h\right]=\sum_{\substack{k=1\\k\neq i,h}}^{M}\frac{\psi_k}{N}\left(1+\frac{\varrho_k\varrho_h}{N-1}\right)+\frac{\psi_i}{N}\left(1+\frac{\varrho_i\varrho_h}{N-1}\right)>\sum_{\substack{k=1\\k\neq i,h}}^{M}\frac{\psi_k}{N}\left(1+\frac{\varrho_k\varrho_i}{N-1}\right)+\frac{\psi_h}{N}\left(1+\frac{\varrho_i\varrho_h}{N-1}\right)=\mathbb{E}\left[\Delta_i\right]\ . \tag{31}$$

APPENDIX G
PROOF OF THEOREM 7

From (7) and (8), we have:

$$\begin{aligned}\alpha_i-\frac{q_i\gamma_i}{\psi_i}=&\sum_{\substack{k=1\\k\neq i}}^{M}\frac{q_i\psi_k\varrho_k}{N(N-1)}-\sum_{\substack{k\in\mathcal{T}\\k\neq i}}\frac{q_k}{N}\left(1+\frac{\left(\varrho_k-\psi_k+1\right)\left(\varrho_i+q_i\right)}{N-1}\right)-\sum_{\substack{k=1\\k\neq i}}^{M}\frac{q_i\psi_k\varrho_k}{N\psi_i(N-1)}\\&+\sum_{\substack{k\in\mathcal{T}\\k\neq i}}\frac{q_iq_k}{N\psi_i}\left(1+\frac{\left(\varrho_k-\psi_k+1\right)\left(\varrho_i+1\right)}{N-1}\right)\\=&\beta_i+\sum_{\substack{k\notin\mathcal{T}\\k\neq i}}\frac{q_i\psi_k\varrho_k}{N(N-1)}\left(1-\frac{1}{\psi_i}\right)+\sum_{\substack{k\in\mathcal{T}\\k\neq i}}\frac{q_i\psi_k\varrho_k\left(1-\frac{1}{\psi_i}\right)-q_iq_k\left(\varrho_k-\psi_k+1\right)}{N(N-1)}\\&+\sum_{\substack{k\in\mathcal{T}\\k\neq i}}\frac{q_iq_k}{N\psi_i}\left(1+\frac{\left(\varrho_k-\psi_k+1\right)\left(\varrho_i+1\right)}{N-1}\right)\ .\end{aligned} \tag{32}$$

Since $1-\frac{1}{\psi_i} \geq 0$, as long as $\psi_i > 0$, then:

$$\sum_{\substack{k \notin \mathcal{T} \\ k \neq i}} \frac{q_i \psi_k \varrho_k}{N(N-1)} \left(1 - \frac{1}{\psi_i}\right) \geq 0 \; . \tag{33}$$

Since $\varrho_k \geq \psi_k \; \forall \; k \in \mathcal{T}, k \neq i$, we get:

$$\sum_{\substack{k \in \mathcal{T} \\ k \neq i}} \frac{q_i q_k}{N \psi_i} \left(1 + \frac{(\varrho_k - \psi_k + 1)(\varrho_i + 1)}{N-1}\right) \geq 0 \; . \tag{34}$$

For $\psi_i > 1$ and $\psi_k > 1, \; \forall \; k \in \mathcal{T} \setminus i$, we have $\varrho_k \geq \varrho_k - \psi_k + 1$ and $\psi_k \left(1 - \frac{1}{\psi_i}\right) \geq q_k$. Consequently:

$$\sum_{\substack{k \in \mathcal{T} \\ k \neq i}} \frac{q_i \psi_k \varrho_k \left(1 - \frac{1}{\psi_i}\right) - q_i q_k (\varrho_k - \psi_k + 1)}{N(N-1)} \geq 0 \tag{35}$$

Note that this expression becomes much greater than zero, when all $\psi_k \in \mathcal{T}, k \neq i$ are much greater than one (i.e. when all targeted vertices other than i have the largest Wants sets). Finally, in case $\psi_i = 1$ and $\psi_k > 1, \; \forall \; k \in \mathcal{T} \setminus i$, we have the third summation in (32) greater than the negative term in the second summation in (32). The theorem follows from all these inequalities and (32).

APPENDIX H
PROOF OF LEMMA 1

We will start this proof from the exact vertex degree expression in (12). Ignoring the content of the different sets, we can derive the expression for the expected degree of a vertex of receiver i as follows:

$$\begin{aligned} \mathbb{E}[\Delta_i] = \mathbb{E}[\Delta_{ij}] &= \sum_{\substack{k=1 \\ k \neq i}}^{M} \mathbb{E}[I_{j \in \mathcal{W}_k}] + \mathbb{E}[I_{j \in \mathcal{H}_k}] \, |\mathcal{W}_k| - \mathbb{E}[I_{j \in \mathcal{H}_k} \cdot |\mathcal{W}_k \cap \mathcal{W}_i|] \\ &= \sum_{\substack{k=1 \\ k \neq i}}^{M} \frac{\psi_k}{N} + \frac{\varrho_k \psi_k}{N} - \mathbb{E}[I_{j \in \mathcal{H}_k} \cdot |\mathcal{W}_k \cap \mathcal{W}_i|] \; . \end{aligned} \tag{36}$$

Substituting (20) in (36), we get:

$$\mathbb{E}[\Delta_i] = \sum_{\substack{k=1 \\ k \neq i}}^{M} \frac{\psi_k}{N} \left[1 + \varrho_k \left(1 - \frac{\psi_i - 1}{N-1}\right)\right] = \sum_{\substack{k=1 \\ k \neq i}}^{M} \frac{\psi_k}{N} \left(1 + \frac{\varrho_k \varrho_i}{N-1}\right) \; . \tag{37}$$

APPENDIX I
PROOF OF LEMMA 2

When the maximal clique κ is chosen for transmission at time t, each member k of the targeted receiver set $\mathcal{T}$ may receive the coded packet with probability q_k. Let X_k be the random variable representing the reception of receiver $k \in \mathcal{T}$ at time t and $\mathbf{X}$ as the random vector of all such random variables. From Lemma 1, we can derive the expression of the expected degree of receiver $i \in \mathcal{T}$ at time $t+1$, conditioned on the random vector $\mathbf{X}$, as follows:

$$\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right] = \sum_{\substack{k\in\mathcal{T}\\k\neq i}} \frac{\psi_k - X_k}{N}\left(1 + \frac{(\varrho_k + X_k)(\varrho_i + X_i)}{N-1}\right) + \sum_{k\notin\mathcal{T}} \frac{\psi_k}{N}\left(1 + \frac{\varrho_k(\varrho_i + X_i)}{N-1}\right) \tag{38}$$

$$= \mathbb{E}\left[\Delta_i^{(t)}\right] + \sum_{\substack{k=1\\k\neq i}}^{M} \frac{\psi_k\varrho_k X_i}{N(N-1)} - \sum_{\substack{k\in\mathcal{T}\\k\neq i}} \frac{X_k}{N}\left(1 + \frac{(\varrho_k - \psi_k + X_k)(\varrho_i + X_i)}{N-1}\right) . \tag{39}$$

Now, we can derive the expected degree of receiver i after serving the maximal clique κ as follows:

$$\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\right] = \mathbb{E}_{\mathbf{X}}\left\{\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\middle|\mathbf{X}\right]\right\}$$

$$= \mathbb{E}\left[\Delta_i^{(t)}\right] + \sum_{\substack{k=1\\k\neq i}}^{M} \frac{\psi_k\varrho_k\mathbb{E}_{\mathbf{X}}\{X_i\}}{N(N-1)} - \sum_{\substack{k\in\mathcal{T}\\k\neq i}} \mathbb{E}_{\mathbf{X}}\left\{\frac{X_k}{N} + \frac{(X_k(\varrho_k - \psi_k) + X_k^2)(\varrho_i + X_i)}{N(N-1)}\right\} . \tag{40}$$

Using the definition of ξ_k and $\Phi_k(x)$ in (8), respectively, we get:

$$\mathbb{E}\left[\Delta_{i\in\mathcal{T}}^{(t+1)}\right] = \mathbb{E}\left[\Delta_i^{(t)}\right] + \sum_{\substack{k=1\\k\neq i}}^{M} q_i\xi_k - \sum_{\substack{k\in\mathcal{T}\\k\neq i}} \Phi_k(q_i) . \tag{41}$$

The expression for $\mathbb{E}\left[\Delta_{i\notin\mathcal{T}}^{(t+1)}\right]$ can be similarly derived using a similar approach.

REFERENCES

[1] R. Ahlswede, N. Cai, S. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, Jul. 2000.

[2] J. Sundararajan, D. Shah, and M. Medard, "Online network coding for optimal throughput and delay - the three-receiver case," *International Symposium on Information Theory and Its Applications (ISITA'08)*, Dec. 2008.

[3] L. Keller, E. Drinea, and C. Fragouli, "Online broadcasting with network coding," *Fourth Workshop on Network Coding, Theory and Applications (NetCod'08)*, Jan. 2008.

[4] E. Drinea, C. Fragouli, and L. Keller, "Delay with network coding and feedback," *IEEE International Symposium on Information Theory (ISIT'09)*, pp. 844–848, Jun. 2009.

[5] S. El Rouayheb, M. Chaudhry, and A. Sprintson, "On the minimum number of transmissions in single-hop wireless coding networks," *IEEE Information Theory Workshop (ITW'07)*, pp. 120–125, Sep. 2007.

[6] A. Blasiak, R. Kleinberg, and E. Lubetzky, "Index coding via linear programming," 2011. [Online]. Available: http://arxiv.org/abs/1004.1379

[7] P. Sadeghi, R. Shams, and D. Traskov, "An optimal adaptive network coding scheme for minimizing decoding delay in broadcast erasure channels," *EURASIP Journal of Wireless Communations and Networking*, Apr. 2010.

[8] D. Nguyen, T. Nguyen, and X. Yang, "Multimedia wireless transmission with network coding," *Packet Video Workshop (PV'07)*, pp. 326–335, Nov. 2007.

[9] H. Seferoglu and A. Markopoulou, "Video-aware opportunistic network coding over wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, pp. 713 –728, Jun. 2009.

[10] ——, "Opportunistic network coding for video streaming over wireless," *Packet Video Workshop (PV'07)*, pp. 191 –200, Nov. 2007.

[11] J. Sundararajan, P. Sadeghi, and M. Médard, "A feedback-based adaptive broadcast coding scheme for reducing in-order delivery delay," *Fifth Workshop on Network Coding, Theory and Applications (NetCod'09)*, Jun. 2009.

[12] P. Sadeghi, D. Traskov, and R. Koetter, "Adaptive network coding for broadcast channels," *Fifth Workshop on Network Coding, Theory and Applications (NetCod'09)*, Jun. 2009.

[13] S. Sorour and S. Valaee, "On minimizing broadcast completion delay for instantly decodable network coding," *IEEE International Conference on Communications (ICC'10)*, May 2010.

[14] ——, "Minimum broadcast decoding delay for generalized instantly decodable network coding," *IEEE Global Telecommunications Conference (GLOBECOM'10)*, 2010.

[15] M. Chaudhry and A. Sprintson, "Efficient algorithms for index coding," *IEEE Conference on Computer Communications Workshops (INFOCOM'08)*, Apr. 2008.

[16] R. Costa, D. Munaretto, J. Widmer, and J. Barros, "Informed network coding for minimum decoding delay," *5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS'08)*, pp. 80 –91, Oct. 2008.